

The 25-year March Toward a Space Physics Data System

Raymond J. Walker
UCLA

Presented at Virtual Observatories in
Space and Solar Physics
Greenbelt, Maryland
October 27, 2004

Why Bother?

*“Since the first satellites had orbited, almost fifty years earlier, trillions and quadrillions of pulses of information had been pouring down from space, to be stored against the day when they might contribute to the advance of knowledge. Only a minute fraction of this raw material would ever be processed; but there was no way of telling what observation some scientist might wish to consult, ten, or fifty, or a hundred years from now. So everything had to be kept on file, stacked in endless air-conditioned galleries, triplicated at the three centers against the possibility of accidental loss. **It was part of the real treasure of mankind, more valuable than all the gold locked uselessly away in bank vaults.**” –*

ARTHUR C. CLARKE, 2001

Space Physics Data Analysis in the 21st Century

- We are entering an era with unprecedented data management challenges.
 - Much more sophisticated instruments are being proposed.
 - Data volumes are growing exponentially. (Future missions will produce $\sim 10^{15}$ B of data!)
 - Studies increasingly require data from more than one source.
 - Data from older missions help us place observations from current missions in context by providing continuity through time.
- Wide spread use of the data will maximize the scientific return of the data!**

The State of Space Data 25-Years Ago

[CODMAC 1 Bernstein et al., 1982]

- “The distribution, storage and communication of data currently limit the efficient extraction of scientific results from space missions.”
- The most successful scientific data activities are run by scientists for scientists.
- There are no technical barriers to impede data handling.
- Recommended a restructuring of the data chain to adhere to basic principles for scientific data management.

Principles for Successful Science Data Management

[*CODMAC Report, Bernstein et al., 1982*]

- **Scientific Involvement**
 - Scientific involvement in all stages of space missions.
- **Scientific Oversight**
 - Peer Review
- **Data Availability**
 - Timely access determined by the scientific community.
 - Easy to use formats. (Make correlative studies easy.)
 - Appropriate ancillary data
 - Enforce contractual obligations for investigators to place data in the archive.
 - Proper documentation.
- **Facilities**
- **Software**
 - Structured, transportable and well documented.
- **Scientific data storage**
 - Stored in a permanent and retrievable form.
- **Data-System Funding**
 - Secure funding from mission overruns.

Since CODMAC 1

- CODMAC 2 and 3– Arvidson et al., 1986; Russell et al. 1988.
 - Geographically distributed, discipline orientated solutions to CODMAC 1 issues.
- CSSP/CSTR Data Panel – Williams, Shea et al., 1984.
 - Focused on accessibility. Create a Solar-Terrestrial Data Catalog.
- **Community Wide Workshop on NASA's Space Physics Data System – Rice University, June 1993.**
- **Space Science Data Systems Technical Working Group**
 - Report of Panel on the State of Space Physics Data - 1997
 - Report of Sun-Earth Connection Study Team (The River Bend Workshop) – 1998.

Community Wide Workshop on NASA's Space Physics Data System

- SPDS was an attempt to carry out the CODMAC recommendations for space physics.
 - Volunteer organization of data users.
 - Small funding to restore in danger of being lost.
 - Did not solve the problem of making well documented data available to the research community.
- Workshop was an attempt to go beyond a small volunteer organization – it was not a success.
- State of Space Physics Data report
 - Importance of research across sub-disciplines not appreciated.
 - PI reluctance to cooperate
 - Incentive is to make discoveries and get them published.
 - Concern others will abuse the data.
 - Resources were inadequate for making the data available.

The River Bend Report

- Plan for a distributed data system for SEC.
- Data managed at a number of sites actively engaged in research.
- Three level 2 entities responsible for managing data at a number of sites.
 - Solar Physics
 - Terrestrial Environment Imaging
 - In situ Space Physics
- Organize views of data in three thematic categories.
 - The Sun as a Star
 - The Sun in Space
 - The Earth in Space
- Science experts organize the data but the data view extends across discipline boundaries.



HEDC - HESSI Experimental Data Center ETH Zurich

- Search RHESSI data products by time
- Go to the main HEDC web interface.
- Anonymous ftp or http access to RHESSI DATA
- Access to RHESSI daily lightcurves
- HESSI Data Analysis Workshop in Zurich (HEDAWZ) was held on April 18, 19, 20 2002 at ETH Zurich, Switzerland
- First Images and Movie

The Ulysses Solar Wind Ion Composition Experiment SWICS

Principal Investigators:

Prof. George Gloeckler
Department of Physics
Institute for Physical
Sciences and Technology
University of Maryland
College Park, MD 20742
USA
Phone (301) 405-6199
Telex (301) 314-9547
E-mail: gg2@umd.edu

Prof. Johannes Oess
International Space
Station Institute
Department of
Physics
University of Maryland
College Park, MD 20742
USA
Phone (301) 405-4892
Telex (301) 314-4892
E-mail: joes@umd.edu

[Description] [User Notes] [References] [WWW Site for SWICS]

Data Available: ~3.5 Hour Averaged Ion Velocity, Temperature, Relative Number Density and Charge States From 341/1990 to 2/15/2003

[Data] [Plots] [Ulysses/SWICS Archive CD-ROM]

ACE SCIENCE CENTER

ACE Home | ACE News | ACE Data | ACE Tools | ACE Help

ACE Science Center
ACE Home | ACE News | ACE Data | ACE Tools | ACE Help

ACE Science Center
ACE Home | ACE News | ACE Data | ACE Tools | ACE Help

ACE Science Center
ACE Home | ACE News | ACE Data | ACE Tools | ACE Help

ELECTRON DRIFT INSTRUMENT (EDI)

EDI Home | EDI News | EDI Data | EDI Tools | EDI Help

EDI Science Center
EDI Home | EDI News | EDI Data | EDI Tools | EDI Help

EDI Science Center
EDI Home | EDI News | EDI Data | EDI Tools | EDI Help

Cluster PEACE High Resolution Data System

Cluster PEACE Home | Cluster PEACE News | Cluster PEACE Data | Cluster PEACE Tools | Cluster PEACE Help

Cluster PEACE Science Center
Cluster PEACE Home | Cluster PEACE News | Cluster PEACE Data | Cluster PEACE Tools | Cluster PEACE Help

TIMED MDC Streamlined Data Product Queries

TIMED MDC Home | TIMED MDC News | TIMED MDC Data | TIMED MDC Tools | TIMED MDC Help

TIMED MDC Science Center
TIMED MDC Home | TIMED MDC News | TIMED MDC Data | TIMED MDC Tools | TIMED MDC Help

The Current Space Physics Data Environment

MACCS Project Information Center
Magnetometer Array for Cusp and Cleft Studies

MACCS Home | MACCS News | MACCS Data | MACCS Tools | MACCS Help

MACCS Science Center
MACCS Home | MACCS News | MACCS Data | MACCS Tools | MACCS Help

Cluster PEACE High Resolution Data System

Cluster PEACE Home | Cluster PEACE News | Cluster PEACE Data | Cluster PEACE Tools | Cluster PEACE Help

Cluster PEACE Science Center
Cluster PEACE Home | Cluster PEACE News | Cluster PEACE Data | Cluster PEACE Tools | Cluster PEACE Help

NASA/ESA Solar and Heliospheric Observatory (SOHO)

SOHO Home | SOHO News | SOHO Data | SOHO Tools | SOHO Help

SOHO Science Center
SOHO Home | SOHO News | SOHO Data | SOHO Tools | SOHO Help

Cluster PEACE High Resolution Data System

Cluster PEACE Home | Cluster PEACE News | Cluster PEACE Data | Cluster PEACE Tools | Cluster PEACE Help

Cluster PEACE Science Center
Cluster PEACE Home | Cluster PEACE News | Cluster PEACE Data | Cluster PEACE Tools | Cluster PEACE Help

FAST High Energy Neutral Analyzer
FAST Science Data Center

FAST Home | FAST News | FAST Data | FAST Tools | FAST Help

FAST Science Center
FAST Home | FAST News | FAST Data | FAST Tools | FAST Help

WIND 3-D Plasma and Energetic Particle Investigation Home Page

WIND Home | WIND News | WIND Data | WIND Tools | WIND Help

WIND Science Center
WIND Home | WIND News | WIND Data | WIND Tools | WIND Help

Transition Region And Coronal Explorer (TRACE)

TRACE Home | TRACE News | TRACE Data | TRACE Tools | TRACE Help

TRACE Science Center
TRACE Home | TRACE News | TRACE Data | TRACE Tools | TRACE Help

Small-Angle Spectrometric Coronagraph (SEC)

SEC Home | SEC News | SEC Data | SEC Tools | SEC Help

SEC Science Center
SEC Home | SEC News | SEC Data | SEC Tools | SEC Help

Small-Angle Spectrometric Coronagraph (SEC)

SEC Home | SEC News | SEC Data | SEC Tools | SEC Help

SEC Science Center
SEC Home | SEC News | SEC Data | SEC Tools | SEC Help

IMAGE High Energy Neutral Analyzer
HENA Science Data Center

IMAGE Home | IMAGE News | IMAGE Data | IMAGE Tools | IMAGE Help

IMAGE Science Center
IMAGE Home | IMAGE News | IMAGE Data | IMAGE Tools | IMAGE Help

State of Solar and Space Physics Data Today

(with help from T. King and S. Joy)

- Carried out a study of sources of space physics data with emphasis on NASA missions.
- Sent a detailed questionnaire to the project scientist of each NASA Sun Earth Connections mission. (All except one responded.)
- Along with a student Erin Means, I examined web pages for all of the missions and most of the instruments – over 200.
- Tested each system by requesting recent data.
- Examined NSSDC holdings for each mission.

Questions

- How hard is it to obtain and use data from current Sun-Earth Connections missions?
 - How easy is it to find the data needed for a given study?
 - If the data were acquired are they available to the scientific community?
 - If the data are available how hard is it to access them?
 - After you have accessed data how hard is it to use them?
 - If you use the data how confident can you be that they are correct?
- Will data be available and useable after the mission ends?
 - Do the mission data repositories have procedures to protect the data from loss?
 - Are the data of archival quality?

The Answers

Data Access and Use

- How easy is it to find the data needed for a given study?
 - *It is not very easy. There are over 200 sources of space physics data. If you aren't one of the cognizanti it is hard to find the data you need. Multi-spacecraft and multi-instrument studies are very difficult.*
- If the data were acquired are they available to the scientific community?
 - *In general they are available. We are much closer to an open data environment than ever before. Missions need to be encouraged to provide all of the data including the highest resolution data.*

More Answers

Data Access and Use

- If the data are available how hard is it to access them?
 - *Mixed. Some missions have excellent interfaces and accessing the data is very easy (e.g. ACE, Geotail). Older missions tended to have the worst interfaces.*
- After you have accessed data how hard is it to use them?
 - *Mixed. Many formats are used. Where documentation is good this is not a serious problem. The quality and thoroughness of the documentation is mixed.*
- If you use the data how confident can you be that they are correct?
 - *Mixed. Many missions either do not have data quality programs or left that question blank.*

More Answers

Data Archiving

- Do the mission data repositories have procedures to protect the data from loss?
 - Yes!
- Are the data of archival quality?
 - No! As noted above the documentation is not uniformly complete. Many data systems have warnings telling users to check with the instrument team before using the data. The instrument team won't always be available. Too many don't get that.

Future Challenges for Space Physics Data Management

- We have moved much closer to an open data environment. We no longer have to convince people of the value of wide access to the data. We have to figure out how to do it in an efficient way.
- The next generation of space physics missions will produce vast data volumes. We will need much better tools for finding, subsetting and accessing the data needed for a given study.
- In space physics in general and the Living with a Star program in particular researchers will increasingly need data which cross traditional discipline boundaries.
- Space physics needs to establish standards and procedures for producing archival quality data.